dr. Mantas Lukoševičius
mantas.lukosevicius@ktu.lt
https://mantas.info/
2022.09.06

ktu
1922
faculty of
informatics
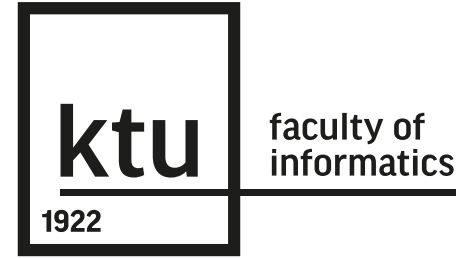
Advanced Machine Learning
P176M010

1. About the Course

Outline

- <mark>About this course</mark>
  - <mark>"Advanced machine learning"?</mark>
  - Goals, structure, etc.
- Getting to know each other
  - About you, your expectations
  - About me, my research projects
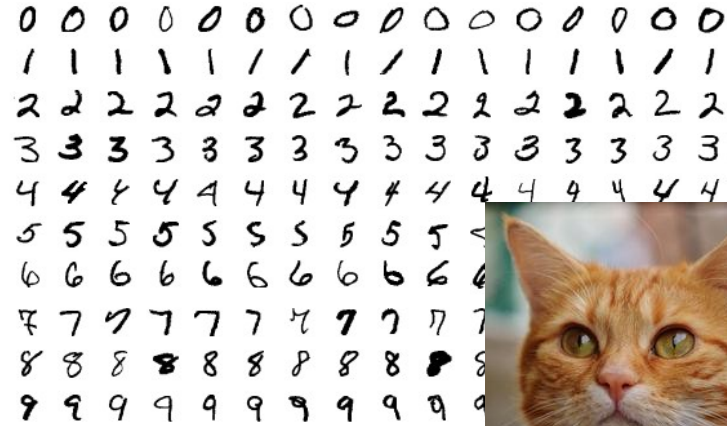- A bit about science and research

"Advanced"?

- As opposed to "basic"

- Not just the basics, a course for Master's students

    – And some *advanced* Bachelor's students ;)

- We will start from the basics but will move on with a tempo

    + links will be provided for a deeper study
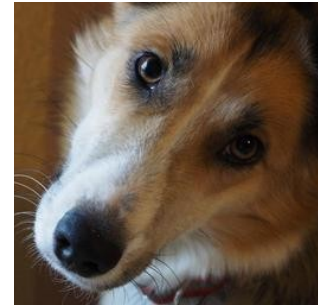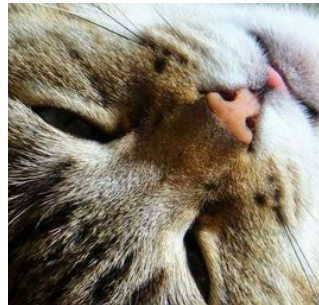
Machine Learning (ML)?

- Most of you know what it is

- Roughly: programs that learn from data, do not need to be explicitly programmed

- Responsible for the recent breakthroughs in Artificial Intelligence

Sample ML problems

Recognize cats and dogs in pictures
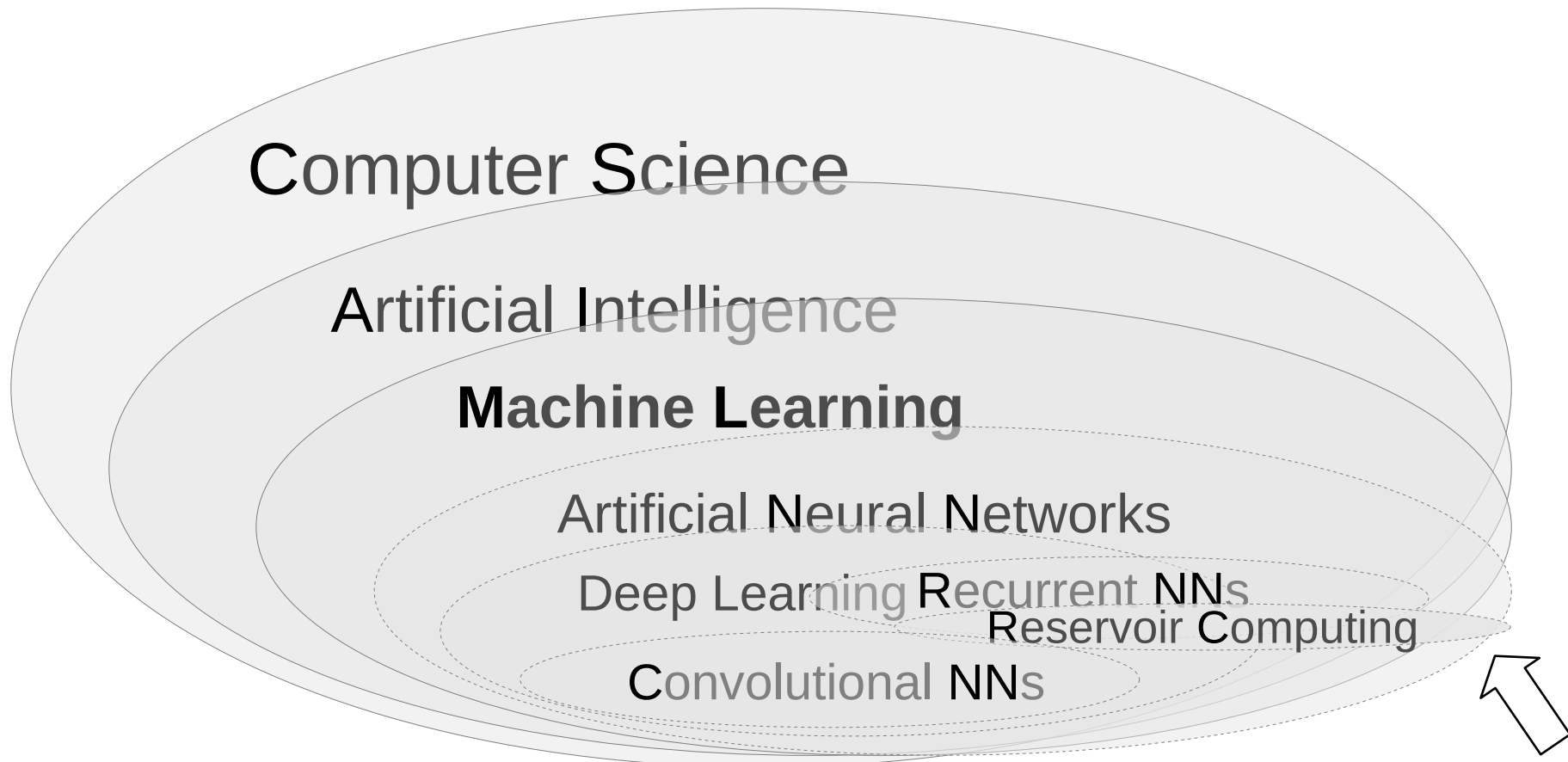
1. About the Course

The Field

Computer Science

Artificial Intelligence

Machine Learning

Artificial Neural Networks

Deep Learning Recurrent NNs

Reservoir Computing

Convolutional NNs

It's actually interdisciplinary!

ML (Deep Learning) dominates modern AI

Mat Velloso
@matvelloso
···

Difference between machine learning and AI:

If it is written in Python, it's probably machine learning

If it is written in PowerPoint, it's probably AI

3:25 AM · Nov 23, 2018 · Twitter Web Client

8,423 Retweets   **889** Quote Tweets   **23.8K** Likes

https://twitter.com/matvelloso/status/1065778379612282885

Why should you care?

- Software is transforming („eating") the world

- AI – even more so

- At the core of modern AI is ML (and DL)

- As future/present IT professionals/leaders you should be at least well familiar with it
  - Not just marketing buzzwords (above), but hands-on

- Not many people are, because it's not trivial

Outline

- About this course
  - "Advanced machine learning"?
  - <mark>Goals, structure, etc.</mark>
- Getting to know each other
  - About you, your expectations
  - About me, my research projects
- A bit about science and research

Prerequisites

At least basics of:

- Programming

- Math

  - Statistics, probability theory, linear algebra, (some calculus?)

- English – *lingua franca* in this field (and others)

1. About the Course

The direct goals of this course

- Be able to apply ML methods to solve real world problems

- To master a "toolbox" of effective ML methods

- Be able to select parameters, train, evaluate ML methods

- To know the classes of ML methods, their principles, pros & cons

- Be able to understand the data, prepare them for ML

- To understand the fundamental ML principles, notions, "laws"

More general goals of this course

- Understanding of the basic AI technologies that are transforming the world

- Understand trends / science / papers in the field better

- New ideas of how to improve ML methods, new applications
  - Better Master's thesis for you
  - Scientific research, startups, etc.

- Helping to build an ecosystem, a community of ML/AI researchers/practitioners
  - AI association of Lithuania – https://lithuania.ai/, https://www.facebook.com/groups/artificialintelligence.lt/, ...

My personal goals with this course

- Stay up-to-date with new trends, have a systematic overview of ML field

- Contribute to the general level of AI/ML understating in this
  (see previous slide)

  - World → Europe → Lithuania → Kaunas → KTU

- Look for / educate new collaborators

  - Maybe future co-workers… (in academia or outside)

- Learn from you

Master's thesis with me?

- To go more in depth with AI/ML

- Directions and topics
  - Some of my research I will briefly present today
  - Individually agreed
  - If you have ideas, let's talk

- Limited capacity…

- Preference to serious(-ly fun) research / application projects

Principles of learning

- We will learn to make machines learn
  - Learn to learn to make machines learn
    - … inception
  - We will learn together. Those who understand better can help the others.
    - Through the official channels: Moodle, Teams, forums – to gain points for active participation
  - There are some similarities between machine and human learning – you will see
- In addition to the theory (Lectures) there will be much practice, individual work and learning
- My teaching is neither a necessary nor a sufficient condition for your learning ;)
- Links for deeper understanding will be all over the slides

Format of the course

- **Lectures** – every even week on Tuesdays 17:30-19:00 online
  - The theory – 8 lectures in total. Video recordings and slides will be made available after each.
- **Labs** – every even week on Tuesdays 19:15-20:45 (or a bit earlier) online
  - 8 practical sessions will help you to apply the theory, master the Python tools needed for mini-projects
- **Mini-projects**
  - 3 individual practical projects, you will have to do it (on time) and present it.
- **Tutorials / seminars** – every odd week on Tuesdays 17:30-19:00 (-20:45) online
  - For different misc. things: discussing mini-projects, have interesting invited speakers, shifted lectures, catching up on things if needed.
  - May be not very regular: no tutorial one week, a double one another week.
- Things might mix up a bit, but we will stay inside the Tuesdays 17:30-20:45 time slots and announce the changes in advance.

Preliminary Curriculum

Theory: lectures

- About this course, science and research
- Machine learning fundamentals and principles
- Data science, preprocessing and feature engineering principles
- Main classical machine learning method classes
- Artificial neural networks
- Deep, convolutional neural networks
- Time series and recurrent neural networks
- Less supervised machine learning, transformer models

Practice: labs, mini-projects; open-source ML/DL technology stack:

- Python
- IDEs, Jupyter, Google Colab
- Visualization
  - Matplotlib, Seaborn, etc.
- NumPy
- Pandas
- SciKit-learn
  - For classical machine learning
- PyTorch
  - For deep learning models
- ...

Grading

- 10% for active participation
  - virtual attendance, informative responses to various activities, feedback, helping colleagues in forums – everything that is not otherwise graded

- 3 x 20% = 60% for 3 mini-projects
  - You will participate in evaluating each other → deadlines
  - You must submit them on time and score ≥45% each
  - Details will be introduced later
  - Possible bonus points for best on-time solutions

- 30% for the final exam
  - Short open questions IRL on a PC, most likely

Ethics

- There are many chances to fool others (and yourself) in Machine Learning (for a short while...)

- An open, friendly, and honest *vibe* is expected

- The goal is to learn: grading, feedback are just the means for the goal

- Use of other people's work should be clearly credited
    - e.g.: "the code is based on https://…", „* X Y helped with..."

Startup-course

Scientific Experiment in Progress

- It's held for the 4th time, 3rd time
  - Online, MS Teams, etc...
  - in English
  - 1st time for 65 (mostly Master's) students
- The material is constantly updated and optimized
  - your feedback is important
- Some technical and organizational hick-ups can be expected, for which I apologize in advance. In case of emergency:
  - don't panic, communicate, be patient, do your best work and you will be fine!

Outline

- About this course
  - "Advanced machine learning"?
  - Goals, structure, etc.
- <mark>Getting to know each other</mark>
  - <mark>About you, your expectations</mark>
  - About me, my research projects
- A bit about science and research

What is your level of English?
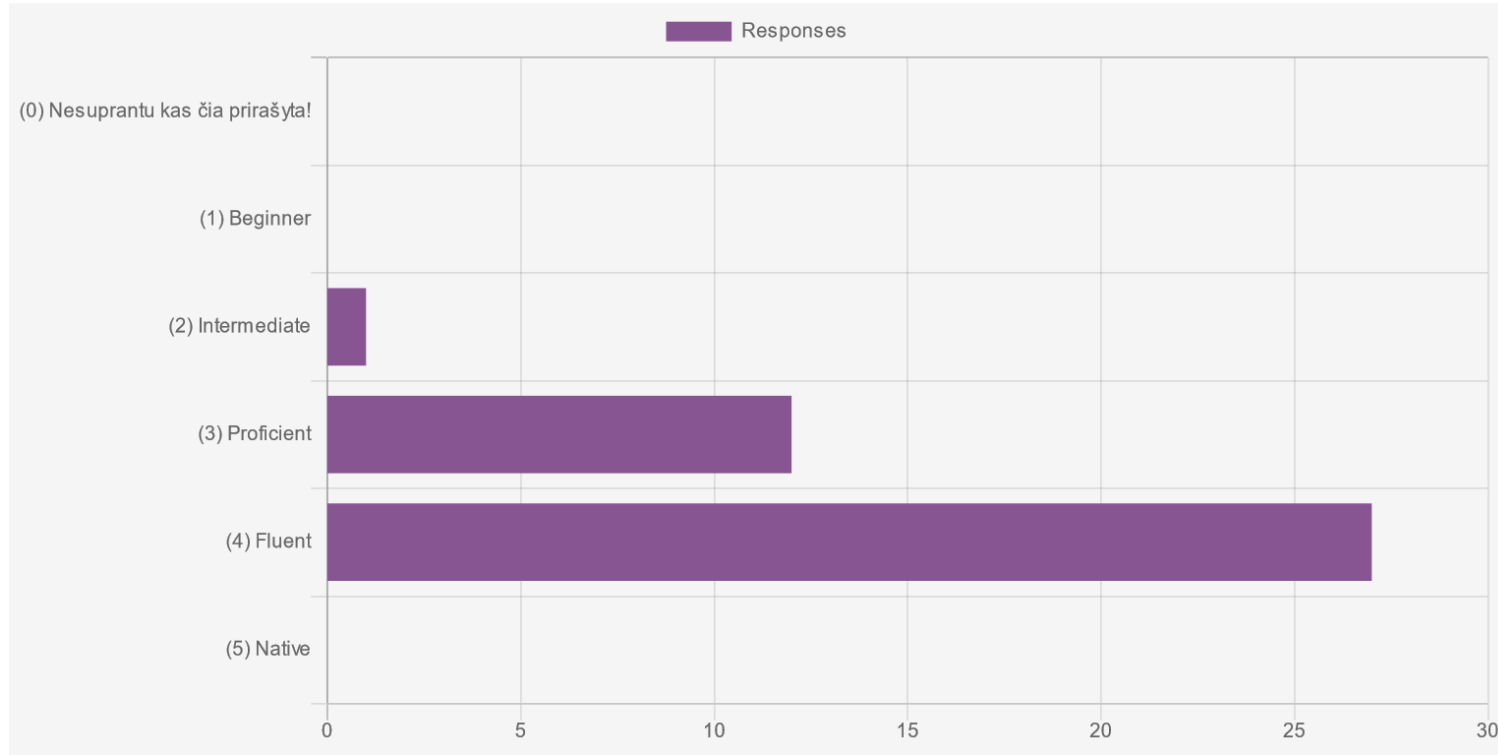
1. About the Course

Where did you study for your Bachelor's?
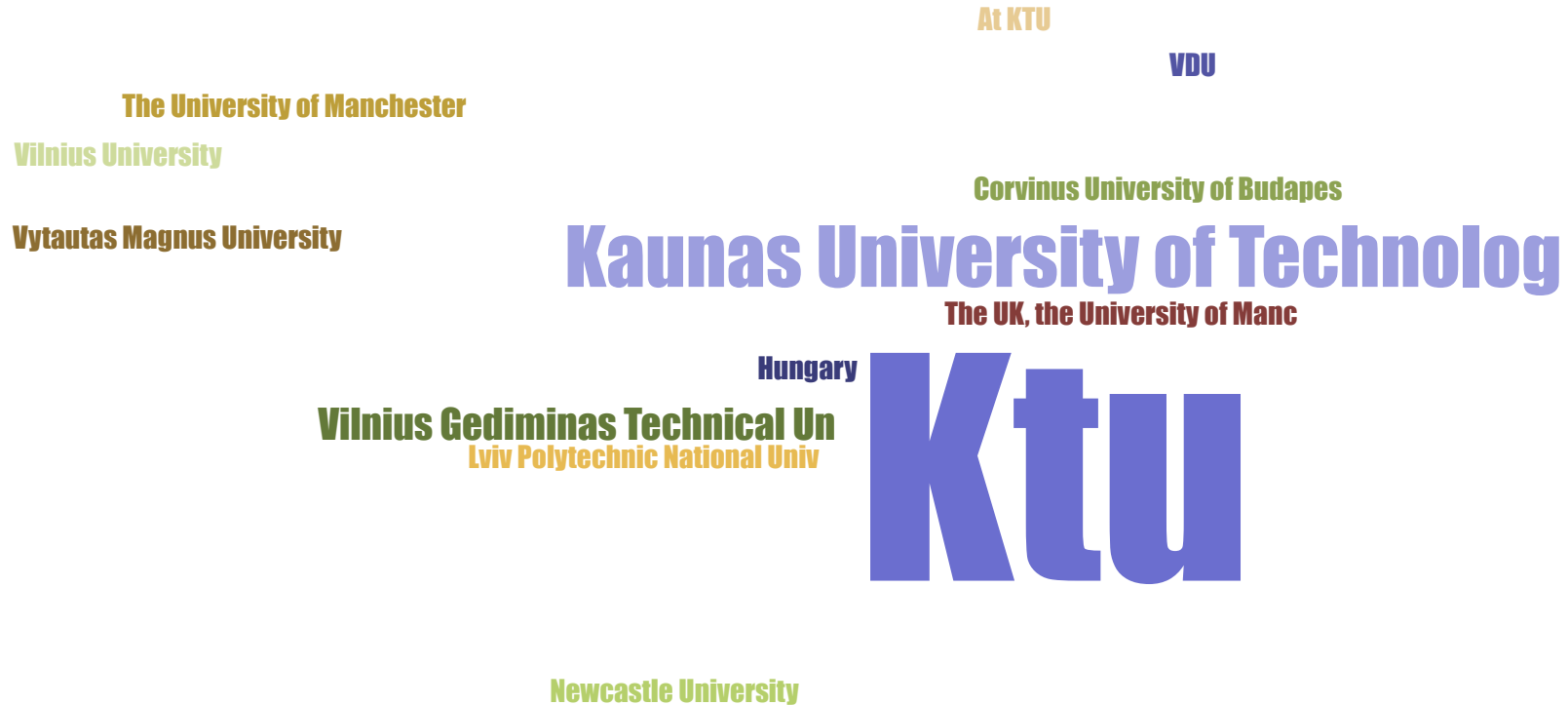
At KTU

VDU

The University of Manchester

Vilnius University

Corvinus University of Budapes

Vytautas Magnus University

Kaunas University of Technolog

The UK, the University of Manc

Hungary

Vilnius Gediminas Technical Un

Lviv Polytechnic National Univ

Ktu

Newcastle University

What did you study for your Bachelor's?

1. About the Course

What is your level of programming experience?

What is your level of programming experience?

1. About the Course
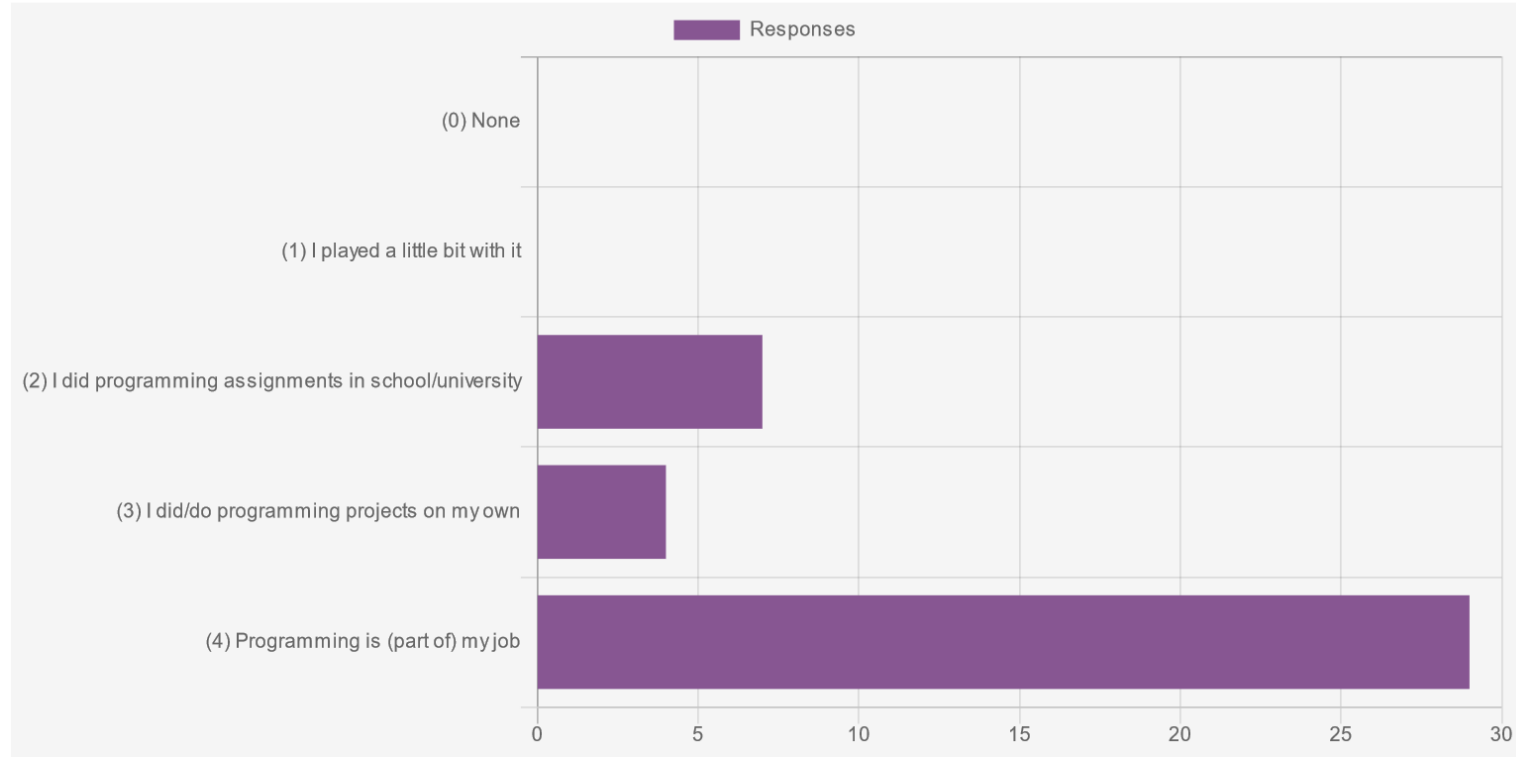
What is your level of Python programming experience?

What is your level of Python programming experience?

1. About the Course
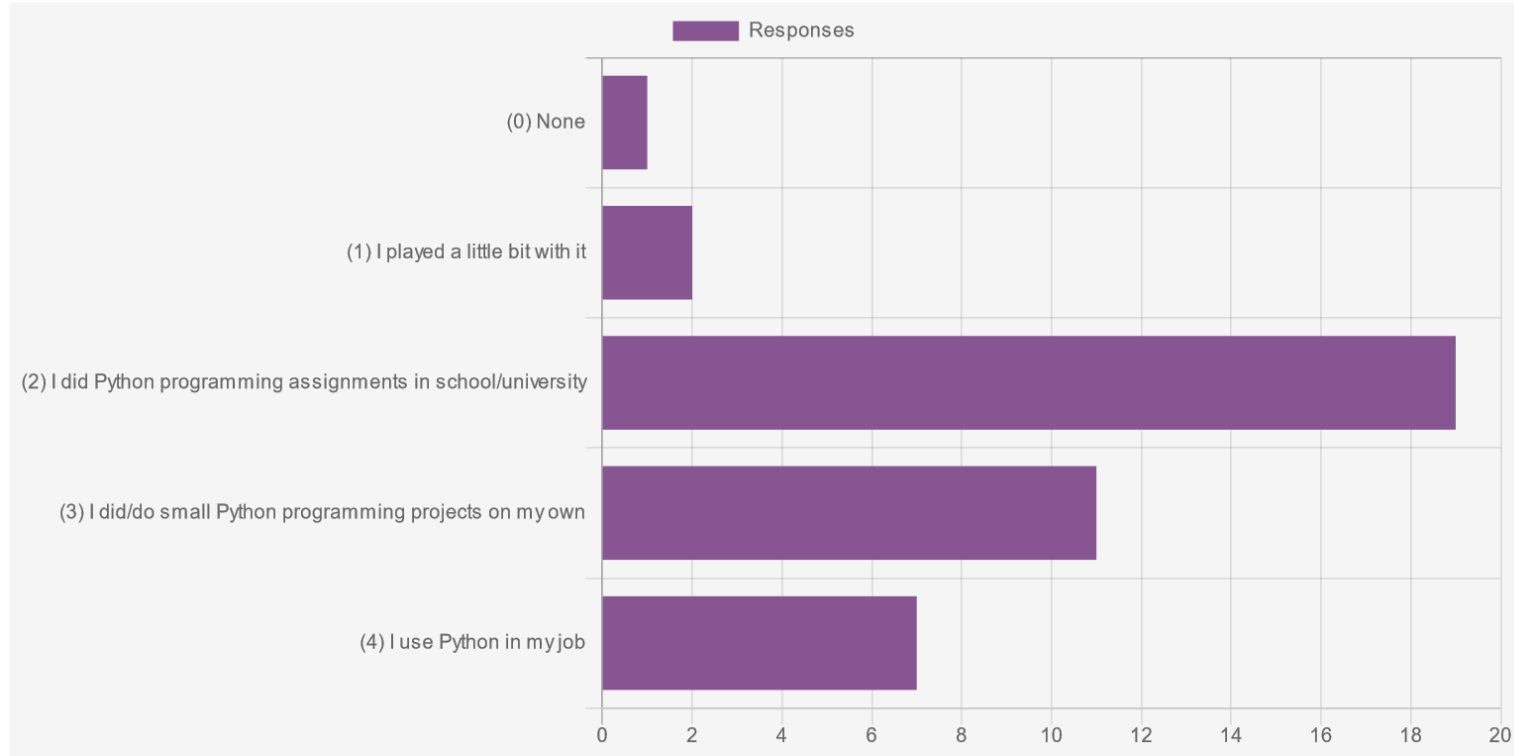
What programming languages do you know, use?*

C

SAS

Carduino JS golang Java

basic

C++ experience

CC++ others

wouldn039t TypeScriptJavaScript SQL much Delphi

Matlab

javascript

JavaScriptTypeScript

Visual CAL

php Python

TS Typescript Scala

*"C" includes C# here

R

Verilog

Kotlin

XAML
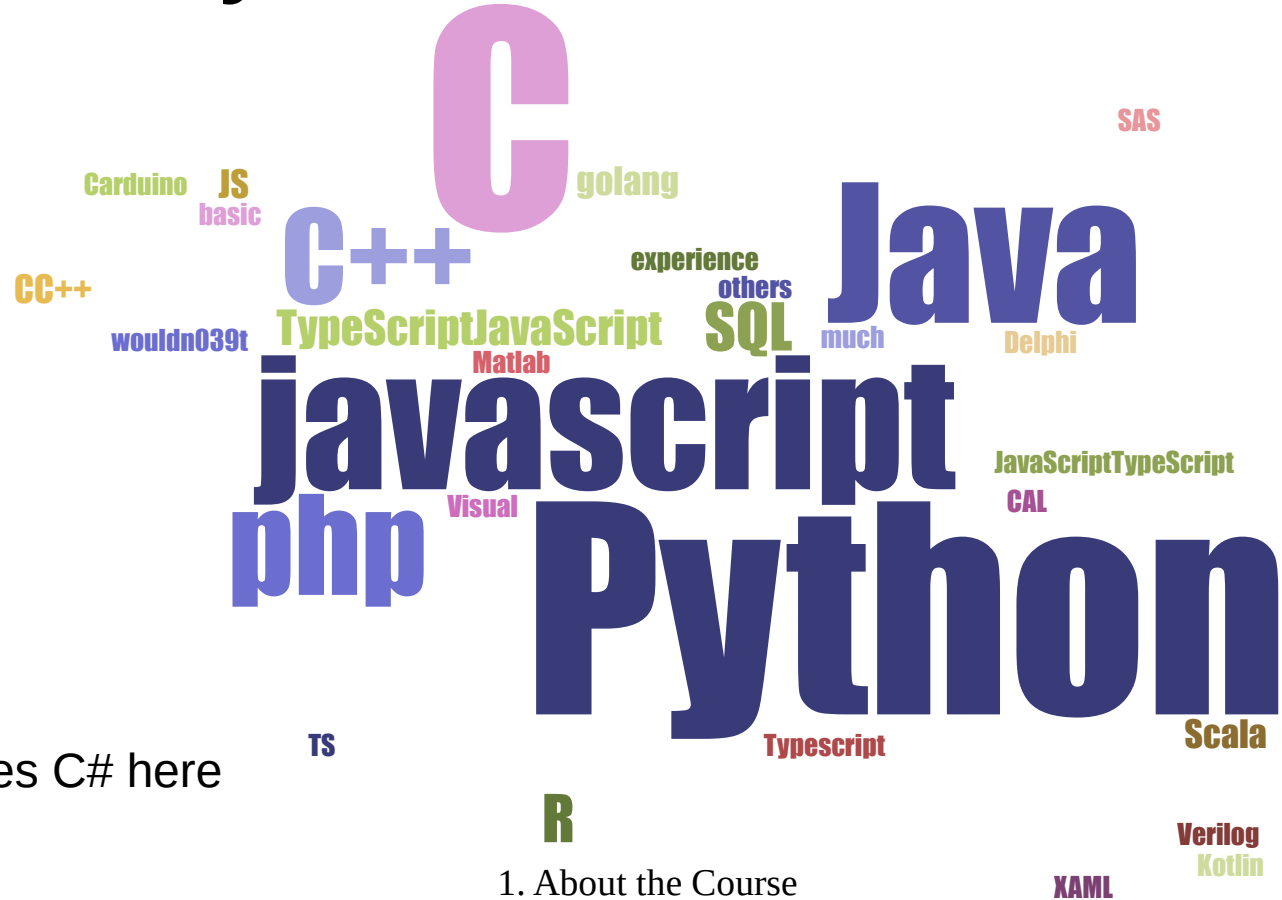
What is your level of ML (/DL) experience?

What is your level of machine learning (/deep learning) experience?

1. About the Course
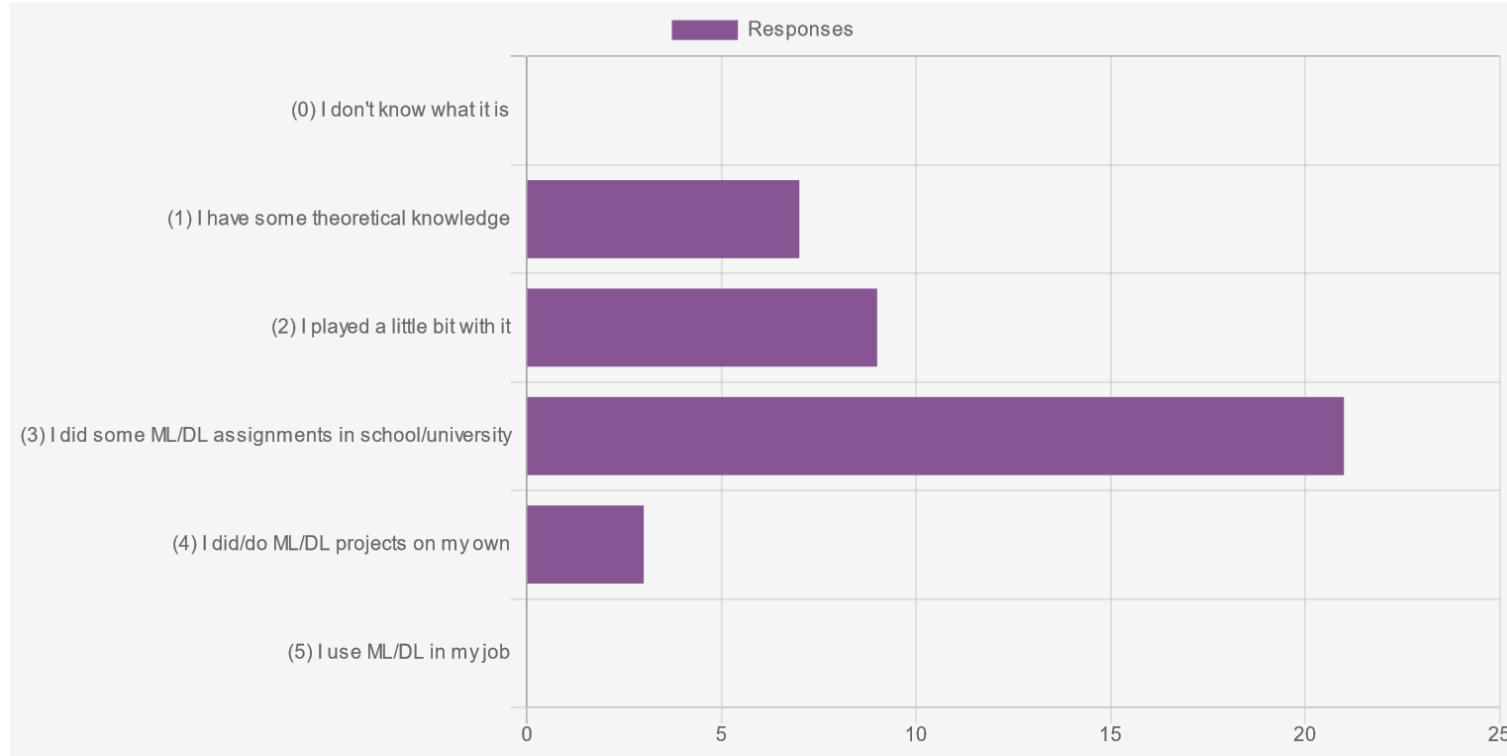
What is your attitude towards mathematics?

What is your attitude towards mathematics?

1. About the Course
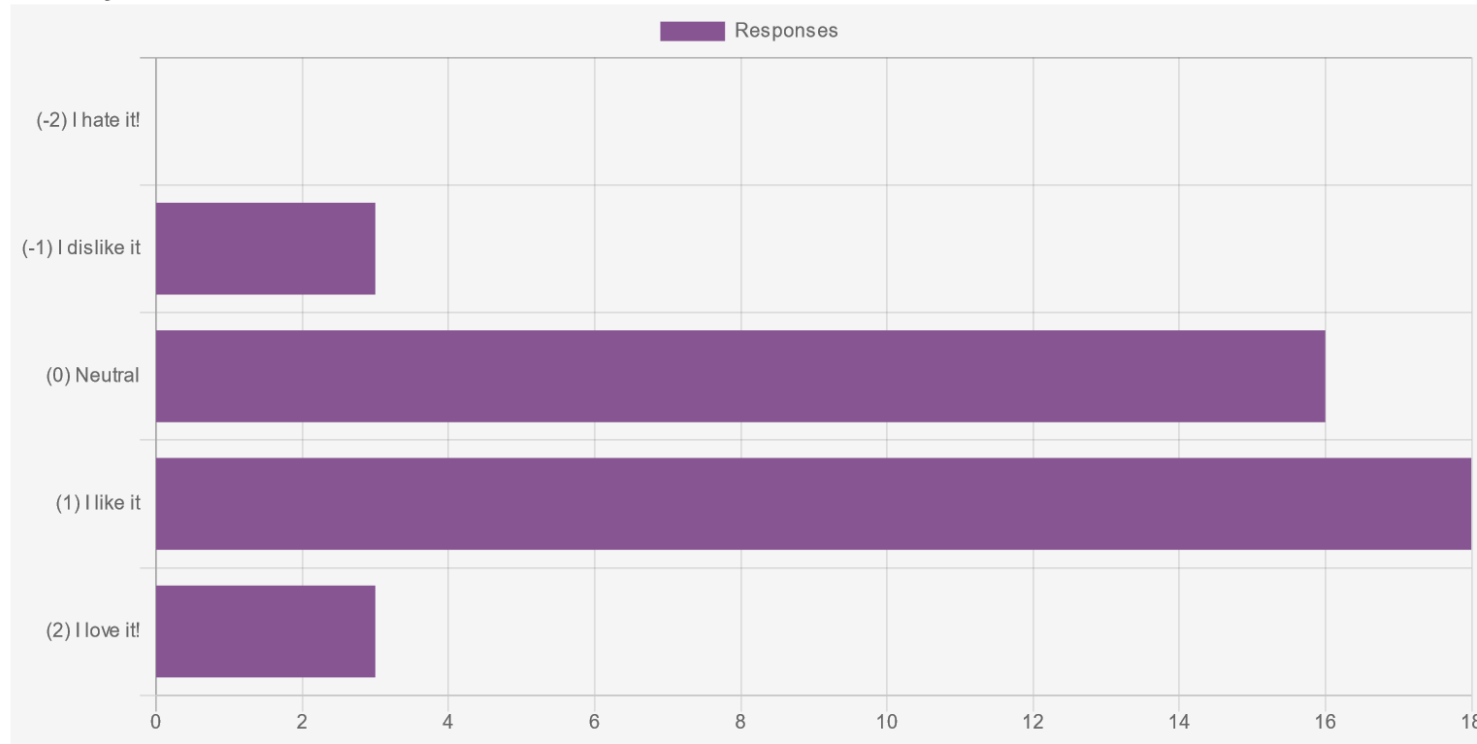
Which of those you could do without Googling or other external help?

Which of those you could do without Googling or other external help? (multiple choice)

1. About the Course
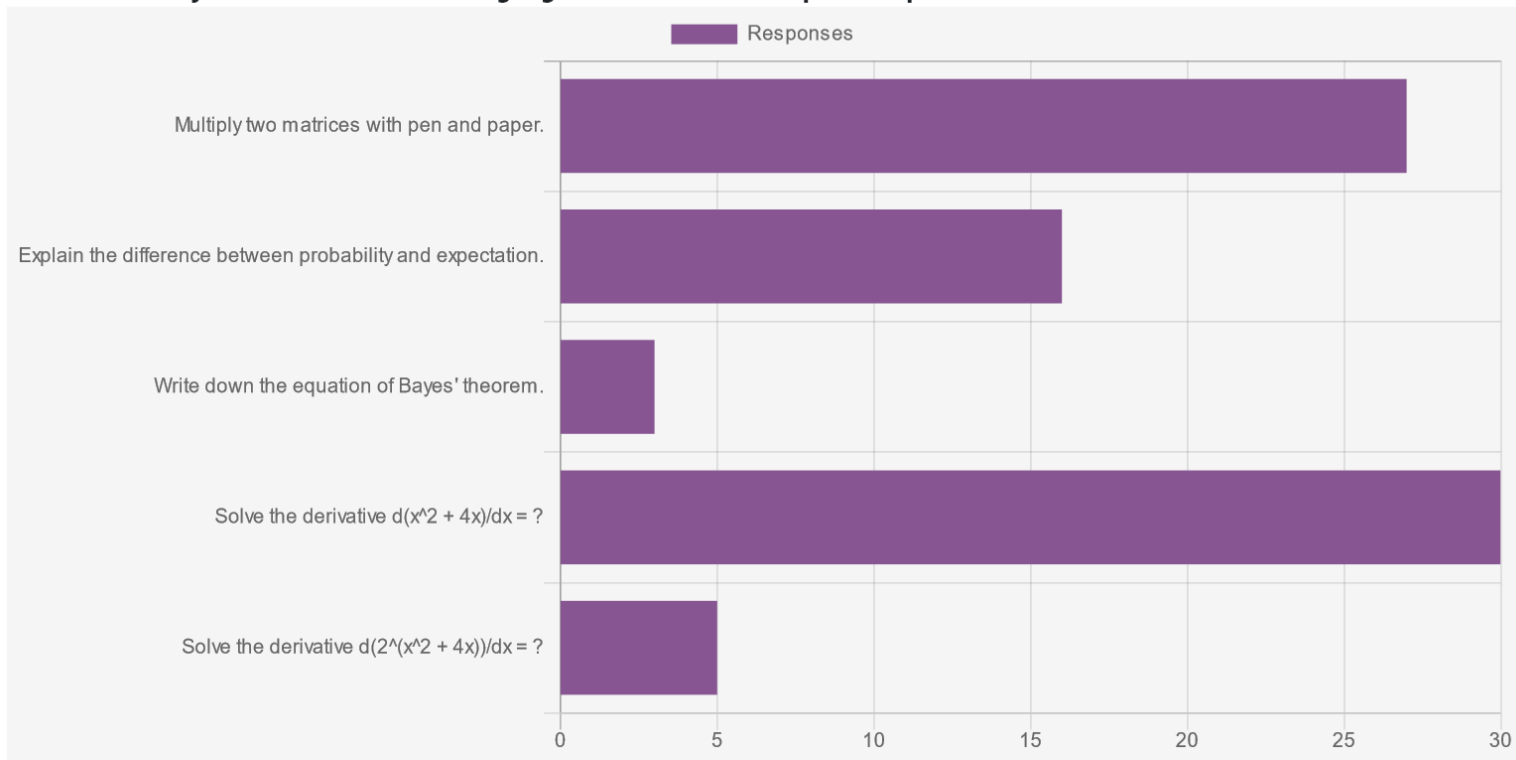
What do you expect to gain/learn in your Master's studies?
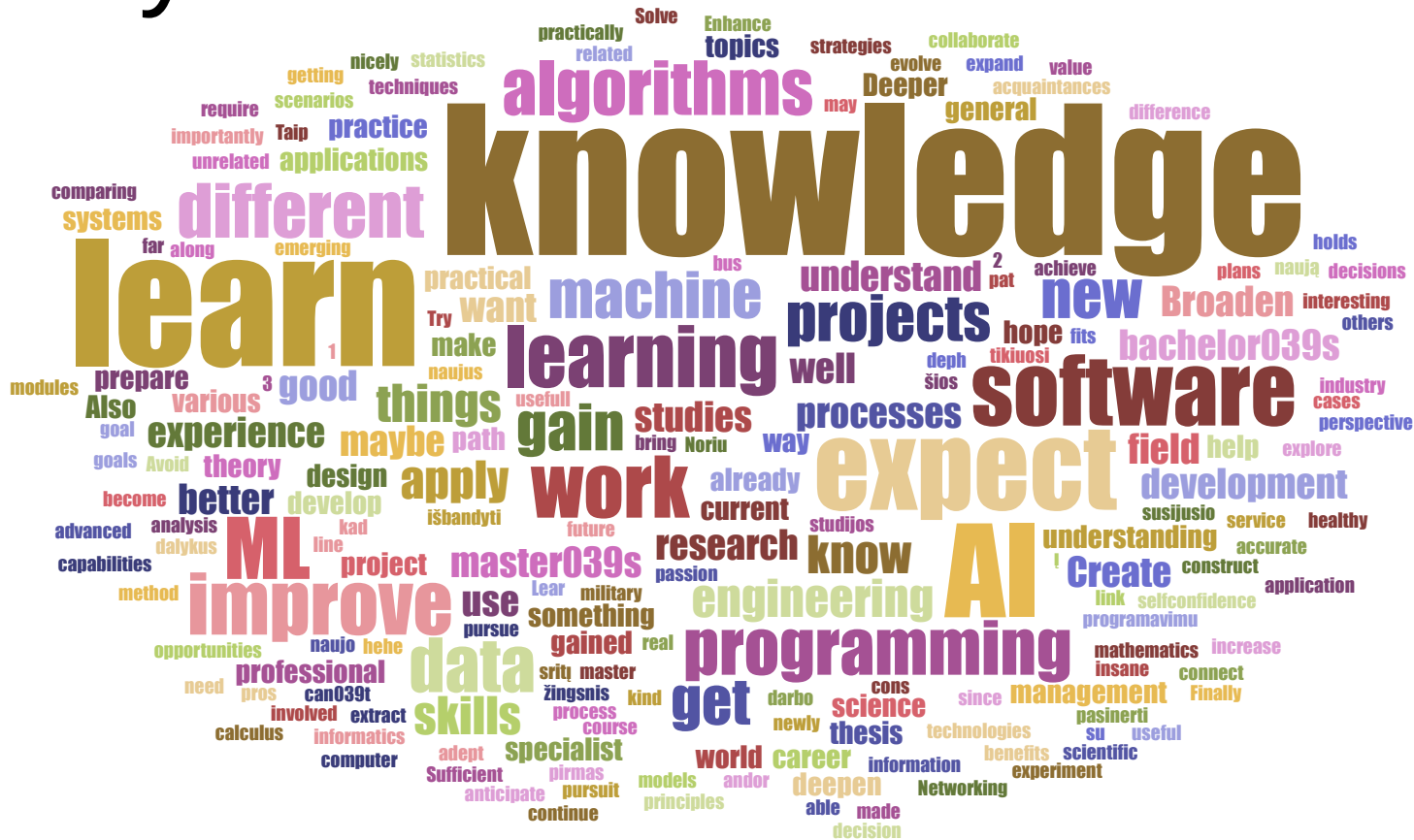
What do you expect to gain/learn in this Advanced ML course?

Make links you have provided available to your classmates?

Would you like to make the links you have provided in your last answer available to your classmates in the AML course and also see theirs? This is a good chance to get to know your virtual classmates better.
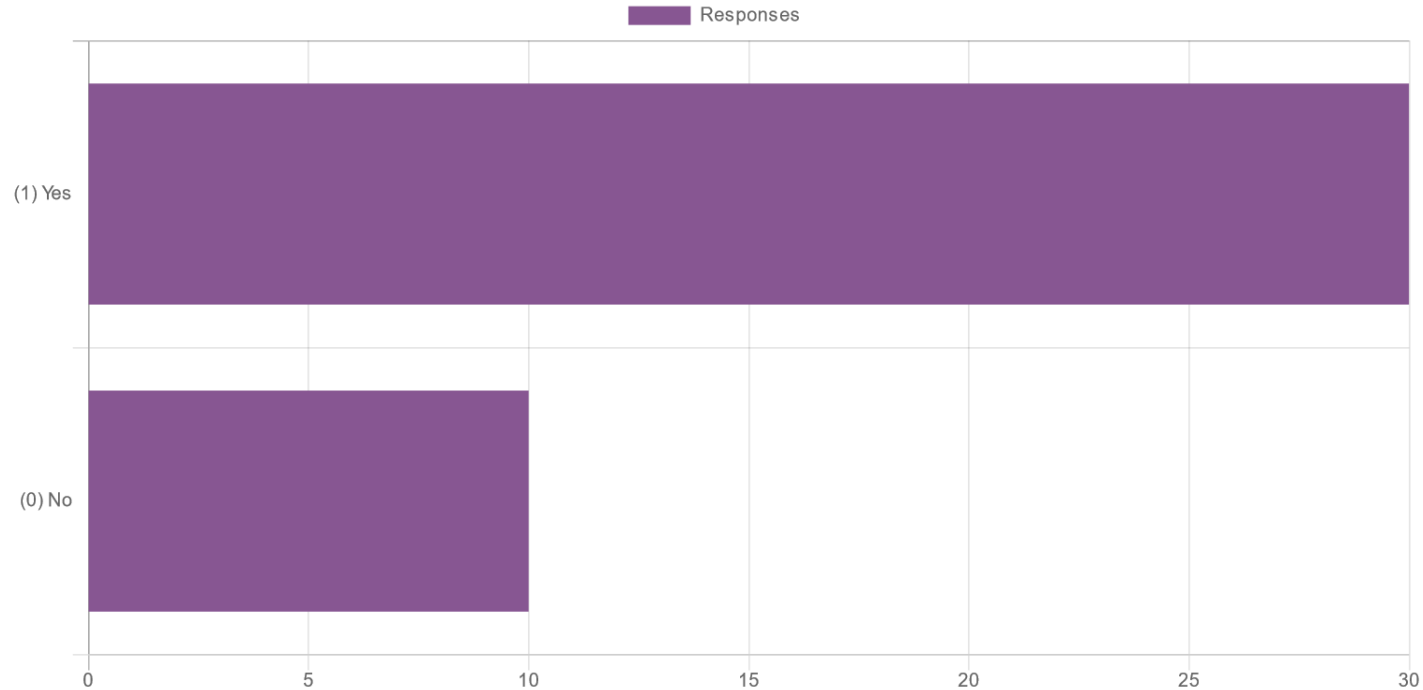
1. About the Course

Outline

- About this course
  - "Advanced machine learning"?
  - Goals, structure, etc.

- Getting to know each other
  - About you, your expectations
  - About me, my research projects

- A bit about science and research

About me

https://mantas.info/

- Started doing ML "before it was cool"

- Education
  - KTU Gymnasium
  - KTU IF Bachelor's
  - Jacobs University Bremen Master's and PhD (ML), Prof. Herbert Jaeger
  - MILA (Université de Montréal) internship, Prof. Yoshua Bengio

- Now
  - Doing ML research
  - Giving this course (obviously)
  - Supervising some students
  - Consulting companies, working in two startups

- A founding member of AI association of Lithuania

- Hobbies:
  - Folk, Lindy Hop, Sci-Fi, different means of transport (sails, bikes, electric unicycles, ...), ...

Some prediction projects in LT

- Ambulance demand predictability
  - Feasibility study
  - Accuracy vs granularity
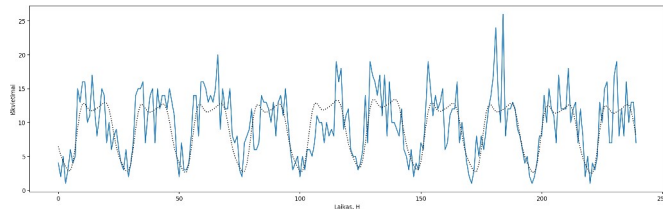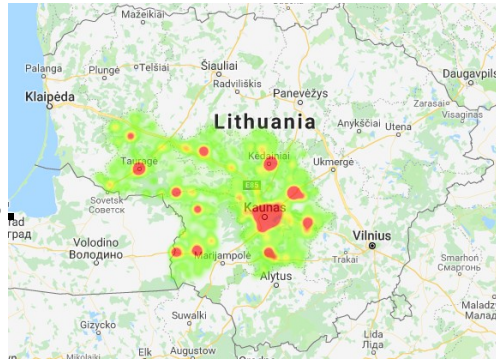
- Nowcasting precipitation with DL
  - Short horizon, first results
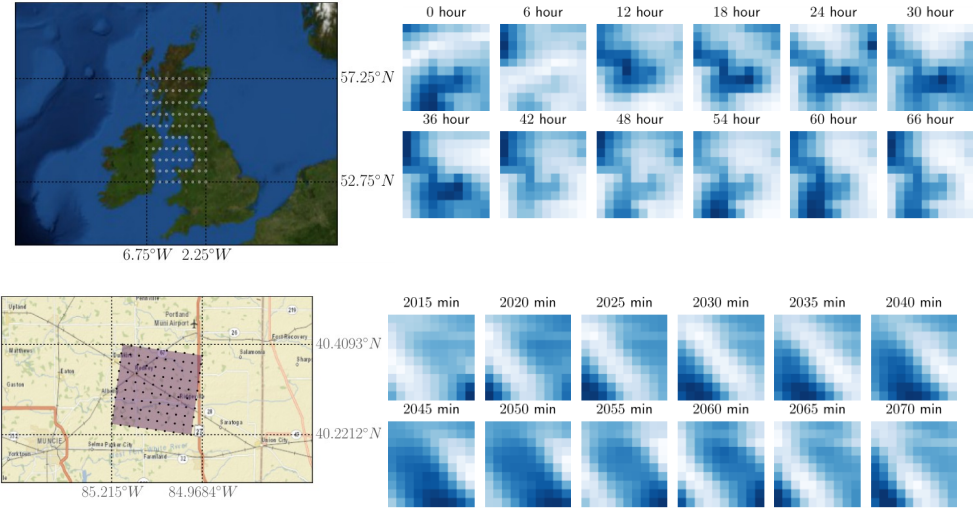   https://kadalis.lt/
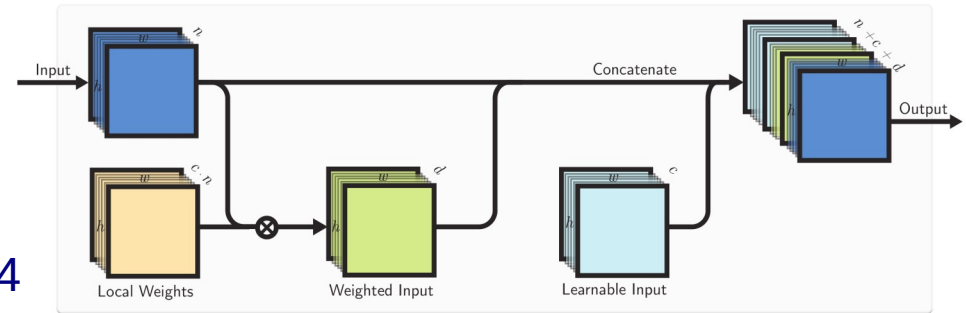 - Radar data
    https://beta.meteo.lt/?pid=radaras

http://ceur-ws.org/Vol-2147/p10.pdf
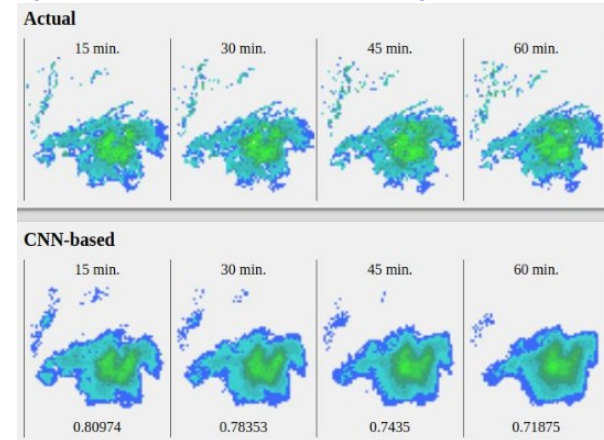
Wind speed predictions

- Better wind speed predictions
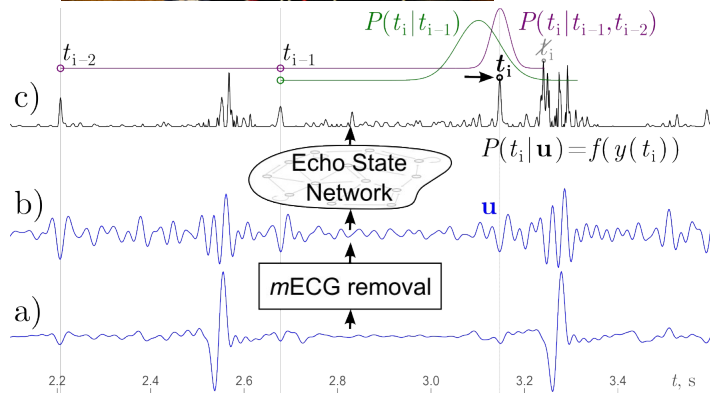 - For renewable energy

- Better „Localized" CNN layers
  https://doi.org/10.3390/en13133440

https://www.youtube.com/watch?v=cie6lCs1nR4
(in Lithuanian)

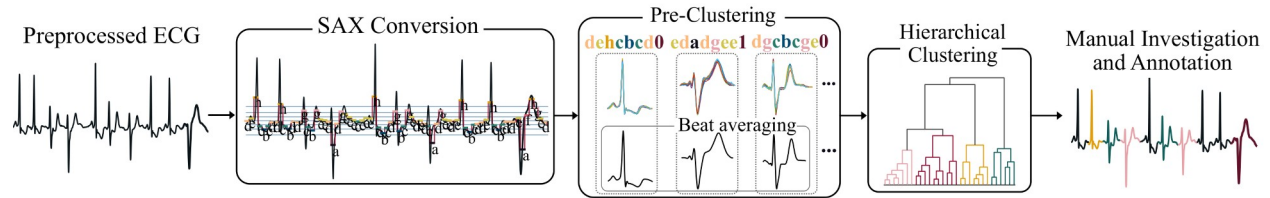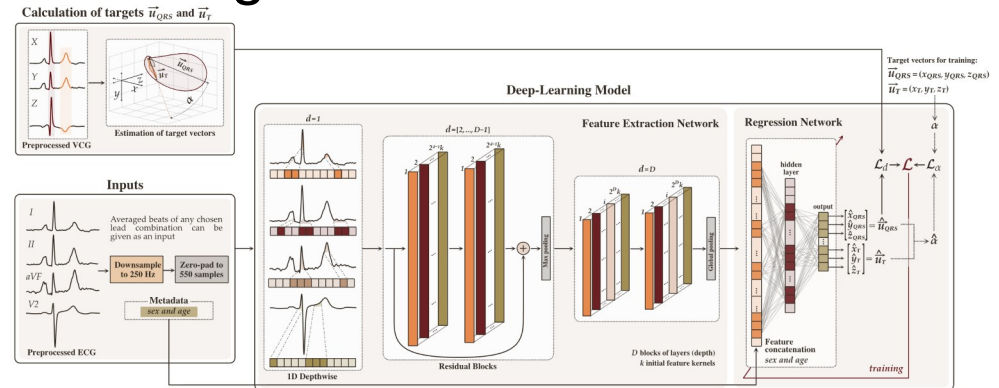# Medical signal applications

- Fetal ECG detection

- ECG clustering in long signals
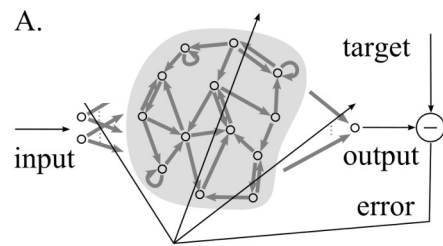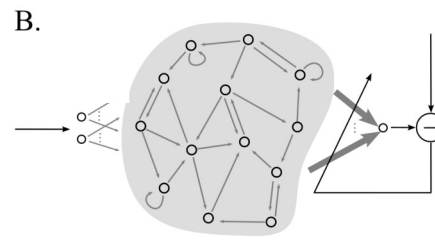
- Detecting heartbeat anomalies with DL

Reservoir computing

- Simpler Recurrent Neural Network (RNN) training

  – E.g., Echo State Network

Backprop-based RNN training   Reservoir Computing

- My overview and tutorial papers are much cited
  https://scholar.google.com/citations?user=kfRwttkAAAAJ

- Cross-validation algorithm for RNNs

  – K-fold C-V of Echo State Networks

  – Almost the same cost as single validation

  – Better generalization

https://arxiv.org/abs/1908.08450

Natural Language Processing

- Lithuanian news
  - Clustering
    - https://arxiv.org/abs/2004.03461
  - Summarization
    - https://arxiv.org/abs/2105.03279
  - Analysis of comments, ...

with PhD student Lukas Stankevičius

- Correcting language errors using deep learning
  - Towards Lithuanian „Grammarly"
    - https://arxiv.org/abs/2203.09963
  - Correcting diacritics and typos in 13 languages
    - https://arxiv.org/abs/2201.13242

Underlying technology: Transformer models

https://mantas.info/science/publications/?tgid=7

Aortic Aneurysm detection startup

- An automatic abdominal aortic aneurysm measurement system from 3D computer tomography scans

   TURIMEDA

  https://www.turimeda.com/

  – In cooperation with Three Thirds
    https://threethirds.ai/cases/turimeda/
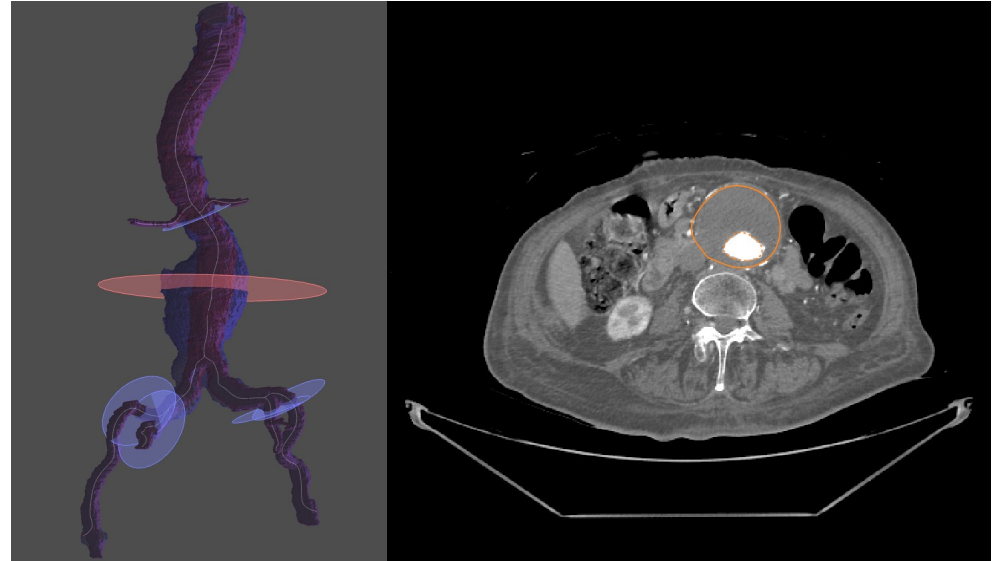
Smart Agrometer startup project

- Tell the missing elements from light specters of plant leaves
  - Precision fertilization

SMART AGROMETER

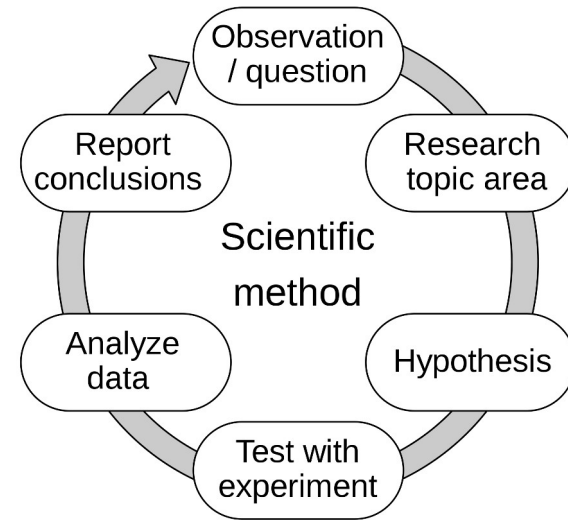https://www.smart.agronom.lt/

Outline

- About this course
  - "Advanced machine learning"?
  - Goals, structure, etc.

- Getting to know each other
  - About you, your expectations
  - About me, my research projects

- A bit about science and research

About Master's studies

- Master's studies already include science

- All KTU IF Master's programs include research

- Research is not necessarily science

    - Research can be: market, client satisfaction, efficiency, technology, etc.

- ML is also very science/research-based

1. About the Course

Research / experiment

- Science is based on experiment
  - You don't know the results in advance
  - The only reliable way to know the nature

- En experiment is performed to confirm or refute a certain hypothesis

- Workflow
  - Get to know the topic
  - Raise a hypothesis / idea / question
  - Implement the experiment to test it
  - Run the experiment, get the data
  - Draw conclusions
  - Repeat...

Observation / question

Research topic area

Report conclusions

Scientific method

Analyze data

Hypothesis

Test with experiment

https://en.wikipedia.org/wiki/Scientific_method

# A classical structure of a research / science paper / thesis

- Abstract
- Introduction – motivation why this is important
- Overview about the field, what related things have already been done by others, how
- Methods – how we propose to solve / research a problem, hypothesis (e.g., "our solution is better") and experiments to test it
  - (Implementation) – level of details enough to reproduce the research
- Results – how good is our solution / how correct is our hypothesis
- Conclusions / discussion / future work
- (Acknowledgments)
- Literature

Standing on the shoulders of Giants

- *„If I have seen further it is by standing on the shoulders of Giants."* – Isaac Newton, 1675

- The amount of scientific knowledge grows exponentially

- New ideas are based on the old ones

- To "enter" the field, know what is there already – usually takes longer than to come up with the new ideas

- http://matt.might.net/articles/phd-school-in-pictures/

Experiments: „*lean*" / „*agile*"

- Scientific principles migrate to business

- Iterative process:
  - Raise a question/hypothesis
  - Create a prototype (MVP)
  - Launch it, get your answer
  - Make conclusions, correct plans, repeat

- The speed is important – „*fail fast*"
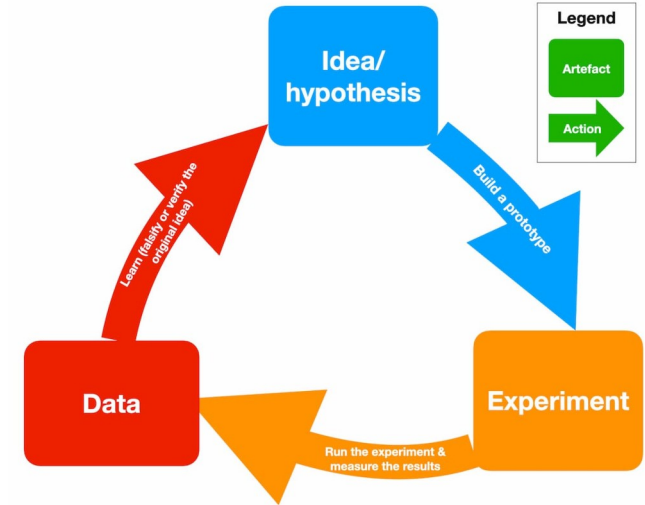
https://en.wikipedia.org/wiki/Lean_startup

Questions?

End of part 1 /8
▮▯▯▯▯▯▯▯

To be continued…

mantas.lukosevicius@ktu.lt

1. About the Course